Article type: Letter

Title: ***sparse panicle1*** **is required for inflorescence development in** ***Setaria viridis*** **and maize**

Author affiliations: **Pu Huang[a,+], Hui Jiang[a,+], Chuanmei Zhu[a], Kerrie Barry[b], Jerry Jenkins[c], Laura Sandor[b], Jeremy Schmutz[b,c], Mathew S. Box[a], Elizabeth A. Kellogg[a], Thomas P. Brutnell[a,*]**

[a] Donald Danforth Plant Science Center, 975 N Warson Rd, St. Louis, MO 63132, USA

[b] Department of Energy Joint Genome Institute, Walnut Creek, California, USA

[c] HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA

[+] These authors contributed equally to this work

[*] *Corresponding author:*

Thomas P. Brutnell

Donald Danforth Plant Science Center, 975 N Warson Rd, St. Louis, MO 63132, USA.

Tel: 1-314-587-1485

Email: tbrutnell@danforthcenter.org

Abstract

Setaria viridis is a rapid life cycle model panicoid grass[1,2]. To identify genes that may contribute to inflorescence architecture and thus have the potential to influence grain yield in related crops such as maize, we conducted an NMU mutagenesis of *S. viridis* and screened for visible inflorescence mutant phenotypes[3]. Of the approximately 2700 M2 families screened we identified four recessive sparse panicle mutants *(spp1-spp4)* characterized by reduced and uneven branching of the inflorescence. To identify the gene underlying the *sparse panicle1* (*spp1*) phenotype, we performed bulked segregant analysis[4] and deep sequencing to fine map it to an approximately 1 Mb interval. Within this interval we identified disruptive mutations in two genes. Complementation tests between *spp1* and *spp3* revealed they were allelic, and deep sequencing of *spp3* identified an independent disruptive mutation in *SvAUX1*, one of the two genes in the ~1Mb interval and the only gene disruption shared between *spp1* and *spp3*. *SvAUX1* was found to affect both inflorescence development and root gravitropism in *S. viridis*. A search for orthologous mutant alleles in maize confirmed a very similar role of *ZmAUX1* in maize, which highlights the utility of *S. viridis* in accelerating functional genomic studies in maize.

Maize (*Zea mays*) is one of the most important crop species globally, and has been used as a genetic system since the early 20th century[5]. Two important goals of maize genetics are to define the mode of action of genes that underlie agronomically important traits and to identify allelic variation that can be exploited in maize breeding programs. Today, emerging food and energy crises demand rapid crop improvement[6], which will require the manipulation of genes underlying important traits. However, the pace of gene discovery in maize is limited by the same traits that make it such a productive crop, namely large stature, complex genome and long life span. Typical forward genetics studies in maize take years from trait discovery to fine mapping of an underlying gene. To date, about 500 genes have been characterized through genetic analysis over decades of study[7]. This contrasts sharply to thousands of genes characterized in *Arabidopsis thaliana*, a rapid cycling eudicot which only became broadly adopted as a model system in the mid 1980s[8,9]. Indeed, many classical inflorescence mutants in maize, such as *barren stalk1*[10] and *barren inflorescence2* (*bif2*)[11] were defined decades before the underlying genes were finally discovered. Although *A. thaliana* is a very successful genetic model, over 140 million years of evolutionary divergence[12] separates the two species that are morphologically, physiologically, and developmentally distinct. For agronomically important traits that do not exist in *A. thaliana*, such as complex inflorescence architecture, C_4 photosynthesis and grain starch accumulation, a grass model is preferred[1].

In recent years *Setaria viridis* has been proposed as a model for food and bioenergy panicoid crops, including maize[1,2]. The lifespan, plant stature and genome size of *S. viridis* are comparable to *A. thaliana*. The extensive gene synteny[13], similar architecture and common habitats that *Setaria* shares with maize suggest its great potential as a translatable model. However, to date neither the utility of *S. viridis* as a genetic model system nor the translatability of discoveries from *S. viridis* to maize have been demonstrated experimentally. In this study, we constructed a mutant population resource for *S. viridis*, and used a forward genetic screen to dissect inflorescence architecture, a complex trait directly related to yield and harvestability in maize and other cereal crops[14,15]. We identified two independent mutations in a single gene, *SvAUX1*, as responsible for major disruptions in inflorescence branch development, which led to sparse panicle (*spp*) phenotypes. We then used a reverse genetics approach to show that a loss-of-function allele in *ZmAUX1*, the maize ortholog of *SvAUX1*, conditions a maize male inflorescence phenotype very similar to *spp1*, namely a reduction in primary branch formation, and thus likely acts through similar genetic mechanisms. Trait discovery to fine map to *SvAUX1* in *S. viridis* took seven months, with greatly lower associated costs and plant growth requirements than an equivalent study would have taken in maize. Importantly, the translatability to maize was readily observed by exploiting the phenotypic similarity and extensive synteny of the two species.

An N-Nitroso-N-methylurea (NMU)-induced mutant population of ~20,000 families was created using reference line A10.1[3]. In total we screened ~2700 mutant families at the M2 generation for phenotypes of interest and identified 61 mutant families for direct sequencing to empirically determine the nature and frequency of chemically induced mutations (Table S1). Most of the changes observed were transition mutations (e.g. G to A single nucleotide polymorphisms (SNPs)) (Fig. S1a). We found a median of 66 homozygous nonsynonymous mutations per mutant family (Fig. S1b). Given this estimation and a Poisson approximation, we estimate that in this population at least 17 mutant alleles exist for 95% of the genes in the genome (Fig. S1c). Thus the *S. viridis* mutant population is a saturated population that can be used for efficient forward and reverse genetics, and multiple alleles of the same gene can be mined.

In an initial screen of ~700 mutant families, we obtained the first allele of the sparse panicle (*spp1*) phenotype (Fig. 1a, i). Defects in *spp1* mutant panicles were detected early in inflorescence development (11 days after germination; A10.1 panicles emerge at approximately 21 days in a chamber under short day condition), as many primary branch primordia failed to initiate, and those that did form were of unequal size. These changes resulted in disrupted phyllotaxis and ultimately the *spp* phenotype (Fig. 1c, d). Direct sequencing of pooled DNA samples from three M3 individuals of this mutant family identified 51 homozygous nonsynonymous mutations (Table S2). We used bulked segregant analysis (BSA) with deep sequencing to fine map the causal mutation[4,16]. To determine the effect of pool size and marker density on mapping resolution and examine the effect of mutant alleles in different genetic backgrounds, five F2 DNA pools from three different crosses were sequenced (Table 1). In all three crosses, F1 plants resembled the A10.1 phenotype (wild-type, WT), and F2 plants displayed a segregation ratio of 3:1 between WT and *spp1* (Table 1). Thus, *spp1* is inferred to be a recessive allele of a single gene. A region of approximately 1Mb near the end of Chromosome 5 showed high homozygosity in all five pools (Fig. 2a; Fig. S2). We did not observe strong differences in mapping resolution in the five pools (Fig. S2) at the same sequencing depth (approx. 30x coverage). Thus, we suggest using a pool size of 30 individuals from a backcross to A10.1 and a sequencing depth of 30x coverage as starting conditions for future mapping attempts. However, if the causal gene is in a region of low recombination, additional individuals may be needed. In this 1Mb region, only two nonsynonymous mutations located in two genes were identified (Table S2). One gene is a ROP interaction partner homolog with unknown function, the other gene is *SvAUX1* (Sevir.5G392400). To validate and fine map the mutation, we identified three other families with *spp* phenotypes in an additional screen of ~2000 mutant families. Among them, *spp3* most closely resembles *spp1* (Fig. 1a). Direct sequencing of *spp3* identified 98 homozygous nonsynonymous mutations in the genome (Table S3). *SvAUX1*, one of the two candidate genes identified in the BSA, is the only gene with a homozygous nonsynonymous mutation in both *spp1* and *spp3* genome-wide (Table S2; Table S3). The F1 hybrid between *spp1* and *spp3* shows a *spp* phenotype, demonstrating *spp1* and *spp3* are non-complementing (Fig. 1b). Phenotypically both *spp1* and *spp3* show decreased plant height, reduced inflorescence branching and spikelet numbers, and increased panicle length compared to A10.1 (Fig. 1i, j; Fig. S3a-c). Collectively, these data indicate that *spp1* and *spp3* are independently mutagenized alleles of *SvAUX1* that condition a *spp* phenotype and will hereafter be referred to as *spp1-1* and *spp1-3*, respectively.

A phylogenetic analysis was performed using the coding sequences of *SvAUX1*, its homologs in other grass species and *A. thaliana* (Fig. 2f). A homolog of *SvAUX1* in *A. thaliana*, *AtAUX1*, is a polytopic membrane protein, and loss-of-function alleles lead to defects in auxin transport[17,18] and agravitropic responses in roots[19]. Loss-of-function alleles of the rice ortholog of *SvAux1*, *OsAUX1*, show agravitropic root growth and defects in lateral root[20] and root hair development[21]. The *SvAUX1* protein is predicted to have 10 or 11 transmembrane domains[17] and likely is necessary for auxin influx[17,18]. The *spp1-1* allele has a premature stop codon (W450*) which truncates the C terminus of the protein (Fig. 2b, d). The C terminus of *AtAUX1* is known to be crucial for its auxin transport function[17,18]. The *spp1-3* allele has a substitution of a charged amino acid for an uncharged one (G332R) in transmembrane domain 7 (Fig. 2b, d), which could lead to conformational changes. Thus, both *spp1-1* and *spp1-3* are likely to be loss-of-function alleles. To test whether *SvAUX1* has a conserved function in root development, we examined root development and gravitropism in *spp1-1* and *spp1-3*. In both mutants, clear agravitropism was observed (Fig. 1e). In a further experiment we changed the direction of the gravity vector during active root growth and agravitropic responses were observed in both primary and lateral roots. Unlike

OsAUX1 which is important for lateral root development, lateral root number was not affected in *spp1-1* mutants (Fig. 1f; Fig. S3d). *SvAUX1* expression is universal in many organs, but its expression in root and panicle is relatively higher than in other organs (Fig. S4c). This corresponds with our observed mutant phenotypes in inflorescence and roots. Through quantitative reverse transcription PCR in emerging panicles of A10.1 and *spp1-1*, we further show that *SvAUX1* is expressed consistently at a lower level in *spp1-1* relative to A10.1 in emerging panicles. Thus, *spp1-1* may encode a message that is both subject to nonsense-mediated decay[22] and generate a truncated protein product (Fig. 1g).

In maize, defects in auxin biosynthesis[23,24], transport[11,25] and signaling[26] have been linked to inflorescence architecture variation, including branching pattern changes. This includes defects in the presumed regulator of auxin effluxer *ZmPIN1*[27,28]. However, to date, the maize auxin influx carrier *ZmAUX1* has not been characterized. In both *AtAUX1* and *OsAUX1* single knockout mutants, no obvious inflorescence phenotype was reported. Only in the triple or quadruple mutant of *AtAUX1* and its three other paralogs were changes of phyllotaxy in shoot and inflorescence meristem observed[29]. Accordingly, the identification of *SvAUX1* as a causal gene for *spp1* hints that auxin influx may regulate inflorescence architecture in maize. Phylogenetic analysis shows *SvAUX1* was co-orthologous to all four homologs in *A. thaliana*. This is in contrast to the *AUX1* homologs of *S. viridis* and maize where clear one-to-one orthology is observed across the phylogeny (Fig. S4). This result highlights the complexity of distant monocot-eudicot comparisons and the ease of using *S. viridis* as model for panicoid crops like maize. We identified a single ortholog (*ZmAUX1*, GRMZM2G127949) of *SvAUX1* in maize through gene synteny[13] and phylogenetic analysis (Fig. 2f, Fig. S4). To examine the potential function of *ZmAUX1*, a *mutator*-tagged allele (*Zmaux1-0*) was obtained from the UniformMu project[30]. The *mutator* insertion is predicted to disrupt the splicing signal of the first intron, and thus is likely to be a loss-of-function allele similar to *spp1-1* and *spp1-3* (Fig. 2c, e). The tassel of *Zmaux1-0* homozygotes shows a clear disrupted branching pattern, similar to *spp1* mutants in Setaria. (Fig. 1h, k). We then backcrossed *Zmaux1-0* to W22 and quantified phenotypes in self pollinated progeny (BC1F2). Homozygous mutant tassels have significantly reduced branch number, spikelet number in the central spike and first primary branch length compared to W22 homozygotes and heterozygotes (Fig. S3e, f, i-l). We also often observed tassel nodes with no primary branches, and some primary branches with very few spikelets in *Zmaux1-0* (Fig. S5). Although the female inflorescence in maize (ear) does not have branches comparable to the tassel or to *S. viridis*, we observed a bald tip on ears in homozygous *Zmaux1-0* and fewer spikelets per row (Fig. S3g; Fig. S5). The disruption of inflorescence development in *Zmaux1-0* is generally less severe than auxin synthesis mutants *sparse inflorescence1* and *vanishing tassel1*[23,24], and auxin transport mutants such as *bif2*[11]. *ZmAUX1* is expressed in many organs, including immature tassel and ear. Furthermore, at least three other homologs are also expressed in these tissues, including immature inflorescences (Fig. S4d). These observations indicate the potential redundancy of gene functions among *AUX1* homologs in maize, and perhaps explains the weaker phenotypes. Also as expected, the *Zmaux1-0* plants displayed a clear agravitropic root growth phenotype (Fig. 1g; Fig. S3h).

Collectively, these results present a compelling case study for gene discovery in maize guided by the genetics of the model grass *S. viridis*. Despite decades of genetic analysis of maize inflorescence development, *ZmAUX1* had not previously been linked to inflorescence architecture, likely because of its subtle phenotype and the lack of clear orthology to *A.thaliana*. In rice, knock-outs of the *ZmAUX1* ortholog *OsAUX1* were recently reported in two separate studies, but both studies lacked descriptions of inflorescence phenotypes[20,21]. This suggests potential functional redundancy of the *Aux1* paralogs in rice. The close phylogenetic relationship of *S. viridis* to maize together with their many anatomical and

biochemical similarities suggests that translating phenotypes from *S. viridis* to maize will be straightforward. Indeed, high throughput suppressor or enhancer screens conducted with the *S. viridis spp1* mutant alleles in greenhouse settings would likely reveal additional components of the auxin influx pathway that would have conserved functions in maize. Unlike maize, *S. viridis* enables rapid gene discovery through forward genetic approaches such as BSA. The time, expense and spatial requirements associated with genetic studies in *S. viridis* are approximately less than 1/3 of that in target panicoid crops like maize, sorghum and switchgrass. Thus *S. viridis* opens the opportunity of gene discovery in panicoid grasses to researchers without access to large field sites or greenhouses. Given the importance of panicoid grasses to global food security and bioenergy production, broader adoption of *S. viridis* by plant scientists throughout the world should rapidly accelerate the understanding of plant gene function and lead to the identification of novel breeding targets to enhance grain yield and biomass production.

References

1. Brutnell, T. P., Bennetzen, J. L. & Vogel, J. P. *Brachypodium distachyon* and *Setaria viridis* : model genetic systems for the grasses. *Annu. Rev. Plant Biol.* **66,** 465–485 (2015).
2. Brutnell, T. P. Model grasses hold key to crop improvement. *Nature Plants* **1,** 15062 (2015).
3. Jiang, H., Huang, P. & Brutnell, T. P. in *Genetics and Genomics of Setaria* (eds. Doust, A. & Diao, X.) **19,** (Springer International Publishing, 2016).
4. Michelmore, R. W., Paran, I. & Kesseli, R. V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 9828–9832 (1991).
5. Emerson, R. A. The inheritance of certain ‘abnormalities’ in maize. *J. Hered.* **8,** 385–399 (1912).
6. Ray, D. K., Mueller, N. D., West, P. C. & Foley, J. A. Yield trends are insufficient to double global crop production by 2050. *PLoS One* **8,** e66428 (2013).
7. Schnable, J. C. & Freeling, M. Genes identified by visible mutant phenotypes show

increased bias toward one of two subgenomes of maize. *PLoS One* **6,** e17855 (2011).

8. Koornneef, M. & van der Veen, J. H. Induction and analysis of gibberellin sensitive mutants in Arabidopsis thaliana (L.) heynh. *Theor. Appl. Genet.* **58,** 257–263 (1980).

9. Redei, G. P. *Arabidopsis* as a genetic tool. *Annu. Rev. Genet.* **9,** 111–127 (1975).

10. Gallavotti, A. *et al.* The role of *barren stalk1* in the architecture of maize. *Nature* **432,** 630–635 (2004).

11. McSteen, P. *et al. barren inflorescence2* Encodes a co-ortholog of the *PINOID* serine/threonine kinase and is required for organogenesis during inflorescence and vegetative development in maize. *Plant Physiol.* **144,** 1000–1011 (2007).

12. Wolfe, K. H., Gouy, M., Yang, Y. W., Sharp, P. M. & Li, W. H. Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc. Natl. Acad. Sci. U. S. A.* **86,** 6201–6205 (1989).

13. Schnable, J. C., Freeling, M. & Lyons, E. Genome-wide analysis of syntenic gene deletion in the grasses. *Genome Biol. Evol.* **4,** 265–277 (2012).

14. Kellogg, E. Floral displays: genetic control of grass inflorescences. *Curr. Opin. Plant Biol.* **10,** 26–31 (2007).

15. Sreenivasulu, N. & Schnurbusch, T. A genetic playground for enhancing grain number in cereals. *Trends Plant Sci.* **17,** 91–101 (2012).

16. Takagi, H. *et al.* MutMap accelerates breeding of a salt-tolerant rice cultivar. *Nat. Biotechnol.* **33,** 445–449 (2015).

17. Swarup, R. *et al.* Structure-function analysis of the presumptive *Arabidopsis* auxin permease *AUX1*. *Plant Cell* **16,** 3069–3083 (2004).

18. Yang, Y., Hammes, U. Z., Taylor, C. G., Schachtman, D. P. & Nielsen, E. High-affinity auxin transport by the *AUX1* influx carrier protein. *Curr. Biol.* **16,** 1123–1127 (2006).

19. Bennett, M. J. *et al. Arabidopsis AUX1* gene: a permease-like regulator of root gravitropism. *Science* **273,** 948–950 (1996).

20. Zhao, H. *et al. OsAUX1* controls lateral root initiation in rice (*Oryza sativa* L.). *Plant Cell Environ.* **38,** 2208–2222 (2015).

21. Yu, C. *et al.* The auxin transporter, *OsAUX1*, is involved in primary root and root hair elongation and in Cd stress responses in rice (*Oryza sativa* L.). *Plant J.* **83,** 818–830 (2015).

22. Baker, K. E. & Parker, R. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr. Opin. Cell Biol.* **16,** 293–299 (2004).

23. Phillips, K. A. *et al. vanishing tassel2* encodes a grass-specific tryptophan aminotransferase required for vegetative and reproductive development in maize. *Plant Cell* **23,** 550–566 (2011).

24. Gallavotti, A. *et al. sparse inflorescence1* encodes a monocot-specific YUCCA-like gene required for vegetative and reproductive development in maize. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 15196–15201 (2008).

25. Gallavotti, A., Yang, Y., Schmidt, R. J. & Jackson, D. The relationship between auxin transport and maize branching. *Plant Physiol.* **147,** 1913–1923 (2008).

26. Skirpan, A., Andrea, S., Xianting, W. & Paula, M. Genetic and physical interaction suggest that *BARREN STALK1* is a target of *BARREN INFLORESCENCE2* in maize inflorescence development. *Plant J.* **55,** 787–797 (2008).

27. Barazesh, S. & McSteen, P. *Barren inflorescence1* functions in organogenesis during vegetative and inflorescence development in maize. *Genetics* **179,** 389–401 (2008).

28. Skirpan, A. *et al. BARREN INFLORESCENCE2* interaction with *ZmPIN1a* suggests a role in auxin transport during maize inflorescence development. *Plant Cell Physiol.* **50,** 652–657 (2009).

29. Bainbridge, K. *et al.* Auxin influx carriers stabilize phyllotactic patterning. *Genes Dev.* **22,** 810–823 (2008).

30. Settles, A. M. *et al.* Sequence-indexed mutations in maize using the UniformMu transposon-tagging population. *BMC Genomics* **8,** 116 (2007).

Online material and methods

NMU-induced S.viridis mutant population

Mutant M1 seeds were obtained after treating A10.1 seeds with NMU (20mM) for 2, 3 and 4 hours[3]. Each M1 plant was self-pollinated to generate a mutant family for screening. To empirically determine mutation rates, DNA samples were extracted from 61 mutant families. For each family, young leaf tissue of one to four M3 or higher generation selfed individuals were collected (Supplementary Table S1). For each family, DNA samples were pooled and sequenced to 20-30X coverage at JGI. Plate-based DNA library preparation for Illumina sequencing was performed on the PerkinElmer Sciclone NGS robotic liquid handling system using Kapa Biosystems library preparation kit. The pool of libraries was then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v4, and Illumina's cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flowcell was performed on the Illumina HiSeq2500 sequencer using HiSeq TruSeq SBS sequencing kits, v4, following a 2x150 indexed run recipe.

All reads were mapped to the *S. viridis* reference genome A10.1 (phytozome v11, http://phytozome.jgi.doe.gov/) using bwa-mem[31]. SNPs were called using GATK unified genotyper[32]. To distinguish SNPs generated by mutagenesis from those caused by residual heterozygosity, only SNPs with alternative alleles occurring in a single family were considered. Effects of SNPs were predicted using snpEff[33]. During SNP calling, we found a large number of SNP calls due to either residual heterozygosity or mapping errors in highly repetitive regions. These SNPs will cause an overestimation of the SNP frequency caused by mutagenesis. To provide a more accurate estimation of the mutation rate, we assume, given a random distribution of mutations caused by NMU, that it is statistically impossible to observe the same SNP more than twice in the 61 sequenced families, whereas for SNPs caused by mapping errors or residual heterozygosity (error-prone SNPs), higher frequencies are expected. Accordingly, we applied a frequency filter to extract NMU-induced SNPs from background. SNPs that occurred in a single mutant family and no other families were considered for frequency calculations. We also generated a list of all the error-prone SNPs that are likely due to mapping error or residual heterozygosity for downstream mapping in this study, as well as future mapping efforts using this NMU population (File S1). All sequencing data were deposited to NCBI-SRA (Table S1).

Direct sequencing and bulked segregant analysis of spp mutants

spp1-1 and *spp1-3* were sequenced directly as part of the panel of NMU mutants. For bulked segregant analysis of *spp1-1*, three crosses were performed between M3 individuals and three different lines: reference line A10.1, another mutant line NMU00290 and a diverse accession TB0155, respectively. We obtained five F2 mutant pools of variable sample sizes (Table 1). A binomial test was applied to examine the fit of mutant and wild-type individuals to a Mendelian ratio of 3:1. Read mapping and SNP calling were conducted as described above. All known error-prone SNPs and SNPs with extremely high coverage (more than 100) were removed from the analysis. The homozygosity frequency of each SNP and a sliding window smooth curve of 10 adjacent SNPs were calculated. Sequences for the five pools were processed separately first, then the resulting allele frequencies of all five pools were merged. For the summarized analysis in order to make the results from different pools comparable, only SNPs overlapping with direct sequencing results from *spp1* were considered.

Plant growth and histology

All plants were grown in a growth chamber under short day conditions (31 °C/22 °C (day/night), 12 hr light/12 hr dark, 50% relative humidity) at the Donald Danforth Plant Science Center. For scanning electron microscope (SEM) work, samples were fixed and dehydrated using standard protocols[34] and critical point dried (CPD) using a Tousimis Samdri-780a. Images were taken with a Hitachi S2600 SEM at Washington University's Central Institute for the Deaf. Input levels and brightness were adjusted in Adobe Photoshop. For the root gravitropism assay, *S. viridis* seeds were sterilized with 20% bleach for 20 min and rinsed 5 times with sterile water. Seeds were kept in water at 4 °C for 3 days and then grown on medium (0.5X Murashige and Skoog basal salt, 1% sucrose, 0.8% agar, pH=5.7). Plants were grown with 200 µmol/m^2s of 12 hour light (6 am-6 pm) , 31 °C day/22 °C night temperature, 50% humidity. Maize seeds were sterilized with 35% hydrogen peroxide for 20 min and rinsed 5 times with sterile water. Seeds were incubated with water in the dark at 30°C for 2 days to germinate and then transferred to medium (1X Hoagland's solution, 0.15% gelzan, pH=6). Plants were grown with 600 µmol/m^2s, 18 hour light (6am-10pm), 28 °C day/ 24 °C night temperature, 50% humidity.

Mutant family UFMU06246, which contains a mutator insertion in GRMZM2G127949 was ordered from Maize Genetics Cooperation (maizecoop.cropsci.uiuc.edu/request/). We identified homozygous individuals through a PCR assay with gene-specific and mutator-specific primers (Table S4). The position of the mutator insertion was confirmed by Sanger sequencing. Seedling gravitropism assays were performed with selfed progeny of these materials. Mutants were backcrossed to W22 and the F1 plants self-pollinated to create a BC1F2 population. Individuals of the BC1F2 population were genotyped using a PCR assay using the same primers that were used previously (Table S4). Plants were propagated under normal greenhouse conditions (14 hour light, 26-28 °C day/22-24 °C night temperature). All panicle and whole plant phenotypes were measured in the BC1F2 population.

Protein structure, phylogenetic and expression characterizations of SvAUX1* and *ZmAUX1

Protter[35] was used to predict transmembrane domains of *SvAUX1* and *ZmAUX1*. Homologs of *SvAUX1* were identified using BLASTN[36] with default settings to primary CDS sequences of grasses and *A. thaliana* (http://phytozome.jgi.doe.gov/). Sequences were aligned using MAFFT[37], and phylogenetic analyses were performed using the maximum likelihood method and GTR+G+I model using RaxML[38]. A rapid bootstrap was performed 1000 times, and the consensus tree was built using the extended majority rule of RaxML. *In silico* data of the expression levels of *SvAUX1* and *ZmAUX1* in *S. viridis* and maize were obtained from the *S. viridis* gene atlas project (http://phytozome.jgi.doe.gov/) and maize gene atlas[39], respectively.

For quantitative reverse transcription PCR, RNA was extracted from emergent panicles of A10.1 and *spp1*, each with three biological replicates. The RNA extraction followed the recommended protocol of TriPure isolation reagent (Sigma-Aldrich). Reverse transcription was performed according to the manufacturer's recommendation (Invitrogen SuperScript III first strand synthesis system), using polyT as primers. To avoid inconsistencies of reference gene expression, we used two reference genes Sevir.9G574400[40] and Sevir.2G354200[41] for comparisons. The primers were designed to be intron-spanning or intron junction spanning to avoid non-specific amplification of genomic DNA (Table S4). Primer efficiencies were determined by a serial 50% dilution using a Roche 480 LightCycler system. Quantitative PCR and data analysis were carried out using recommended settings of the Roche 480 LightCycler system and LightCycler 480 SYBR Green I Master. The normalized relative quantity to the two reference genes was estimated using the method described in[42].

Data availability

All sequencing data can be accessed through NCBI-SRA. Detailed accession numbers were listed in Table S1. All the error-prone SNPs from the NMU population that are likely due to mapping error or residual heterozygosity were listed in File S1.

Author Contributions

P.H., H.J. and T.P.B. conceived and designed the study. H.J. performed the screen, crosses and DNA extraction. P.H. and H.J. performed bulked segregant analysis. K.B., J.J., L.S. and J.S. performed library construction and sequencing. P.H. performed sequencing and other data analysis. P.H., H.J., C.Z. and M.S.B. performed phenotypic characterizations in *S. viridis* and maize. P.H., H.J., C.Z., E.A.K. and T.B. wrote the manuscript.

Acknowledgements

The authors thank Dr. Adam Bray for his help in maize root gravitropism assay, Dr. Christine Shyu for her help in qRT-PCR, and the DDPSC greenhouse staff for plant care. The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the US Department of Energy under contract number DE-AC02-05CH11231. This work was also supported by a Department of Energy grant to T. Brutnell (DE-SC0008769), and a National Science Foundation grant to E. A. Kellogg (IOS-1413824).

Reference for online methods:

31. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **1303.3997,** (2013).
32. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20,** 1297–1303 (2010).
33. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6,** 80–92 (2012).
34. Hodge, J. G. & Kellogg, E. A. Patterns of inflorescence development of three prairie grasses (Andropogoneae, Poaceae). *Int. J. Plant Sci.* **175,** 963–974 (2014).
35. Omasits, U., Ahrens, C. H., Müller, S. & Wollscheid, B. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* **30,** 884–886 (2014).
36. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410 (1990).
37. Katoh, K., Misawa, K., Kuma, K.-I. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30,** 3059–3066 (2002).
38. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30,** 1312–1313 (2014).
39. Stelpflug, S. C. *et al.* An Expanded Maize Gene Expression Atlas based on RNA Sequencing and its Use to Explore Root Development. *Plant Genome* **9,** 0 (2016).
40. Lambret-Frotté, J. *et al.* Validating Internal Control Genes for the Accurate Normalization of qPCR Expression Analysis of the Novel Model Plant Setaria viridis. *PLoS One* **10,** e0135006 (2015).

41. Dong, M. A., Farré, E. M. & Thomashow, M. F. Circadian clock-associated 1 and late elongated hypocotyl regulate expression of the C-repeat binding factor (CBF) pathway in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 7241–7246 (2011).
42. Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* **8,** R19 (2007).

411
412
413 **Table 1.** Populations for bulked segregant analysis

Parent 1	Parent 2	Segregation WT:*spp1*	p [a]	No. of F2 individuals pooled	Median fold coverage (genome wide)
NMU00629.3m (*spp1-1*)	TB0155	147:45	0.54	44	29
NMU00629.3m (*spp1-1*)	NMU00290[b]	304:105	0.67	30	24
				60	30
NMU00629.3m (*spp1-1*)	A10.1[b]	272:94	0.70	30	30
				64	26

414
415 [a] p value of two-tailed binomial test of expected ratio 3:1
416 [b] two DNA pools of different sizes were made from mutants generated by each cross.
417

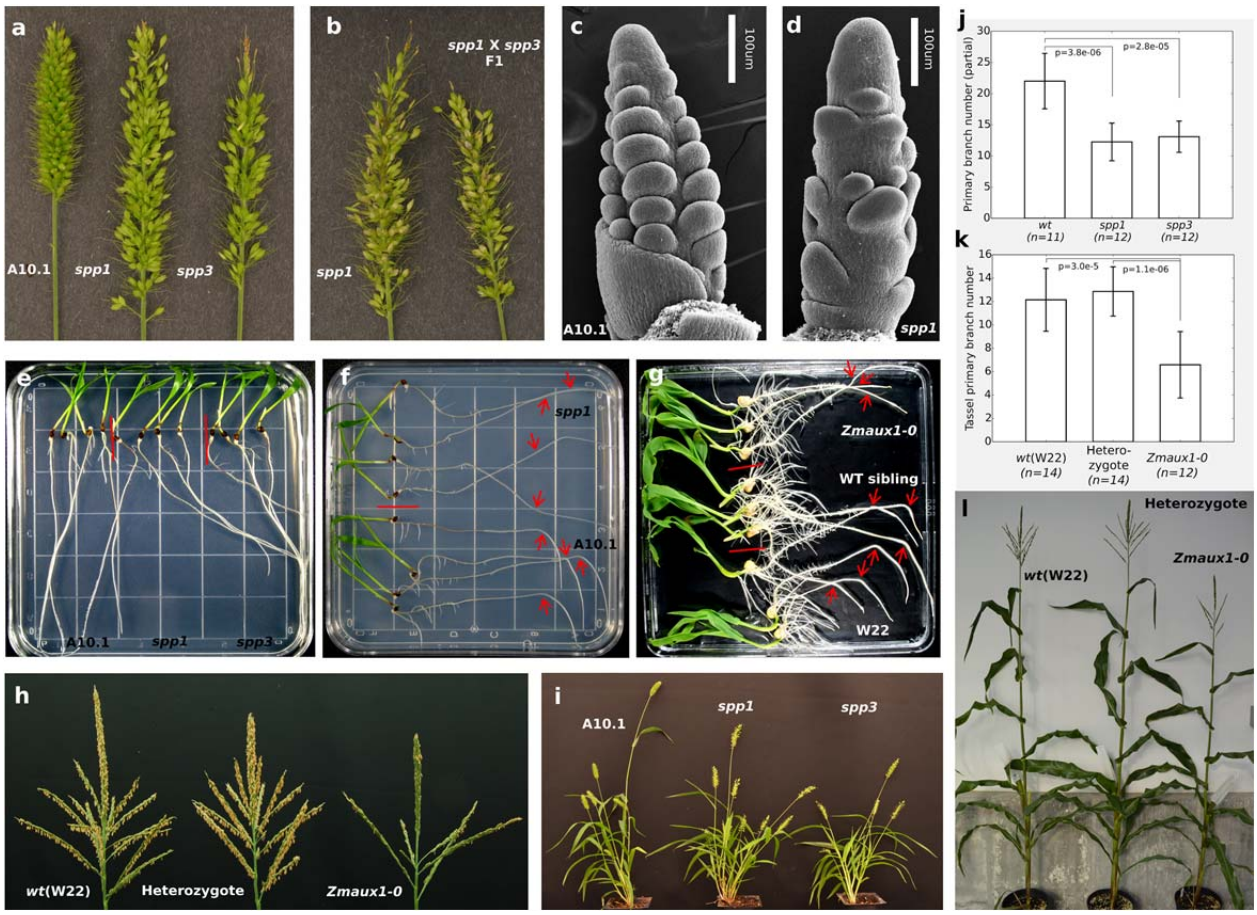


Figure 1. Characterization of *spp1* and *spp3* in *S. viridis* and *Zmaux1-0* in maize. **a)** Reduced branching of panicles in *spp1* and *spp3*. **b)** Complementation test between *spp1* and *spp3*; hybrid shows *spp* phenotype. **c,d)** Early panicle development (11 days after germination) of A10.1 and *spp1*. **e)** Root growth of A10.1, *spp1* and *spp3*, 6 days after germination. **f)** Root gravitropism assay for A10.1 and *spp1*. **g)** Root gravitropism assay for *Zmaux1-0*, segregating wild type siblings from the same mutant family and wild type background (W22). In F and G, the direction of the gravity vector was changed 90° clockwise at 4 days after germination, and the photo was taken at day 5 (*S. viridis*) or day 6 (maize). Red arrows indicate the positions of root tips at day 4 before rotating the plates. Seedlings of different genotypes are indicated by red vertical lines. **h)** Inflorescence phenotype of wild type (W22), heterozygote and homozygote of *Zmaux1-0*. **i)** Mature A10.1, *spp1* and *spp3* in *S. viridis*. **j)** Inflorescence primary branch number of A10.1, *spp1* and *spp3* in *S. viridis*, measured from bottom, 0.5 cm to 1.5 cm. **k)** Tassel primary branch number of WT, heterozygote and homozygous *Zmaux1-0*, from the bottom 5 cm. All bar heights show mean values, error bars show standard deviations, p-values were calculated from two-tailed t-tests and n= number of individuals examined. **l)** Mature plant of W22, heterozygote and *Zmaux1-0*.

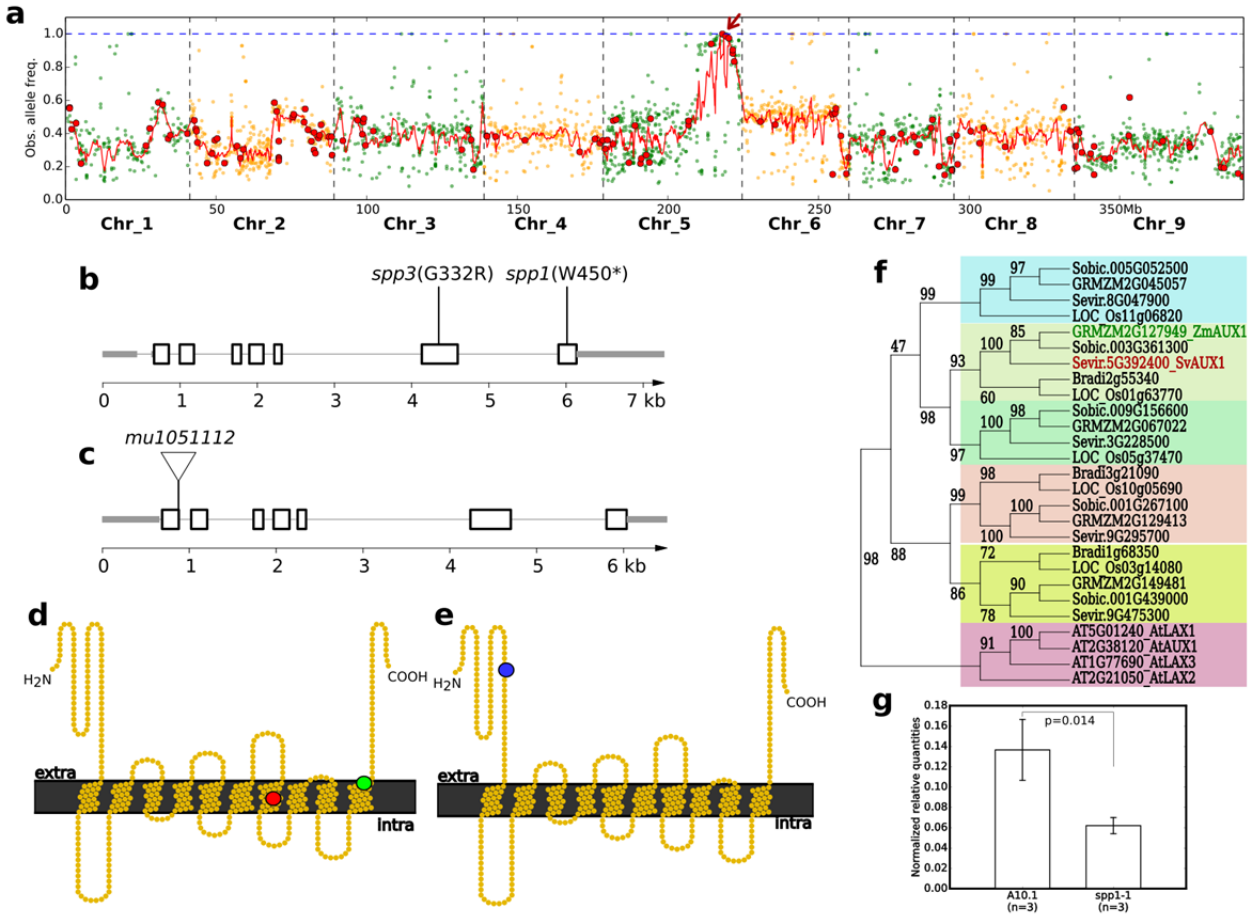
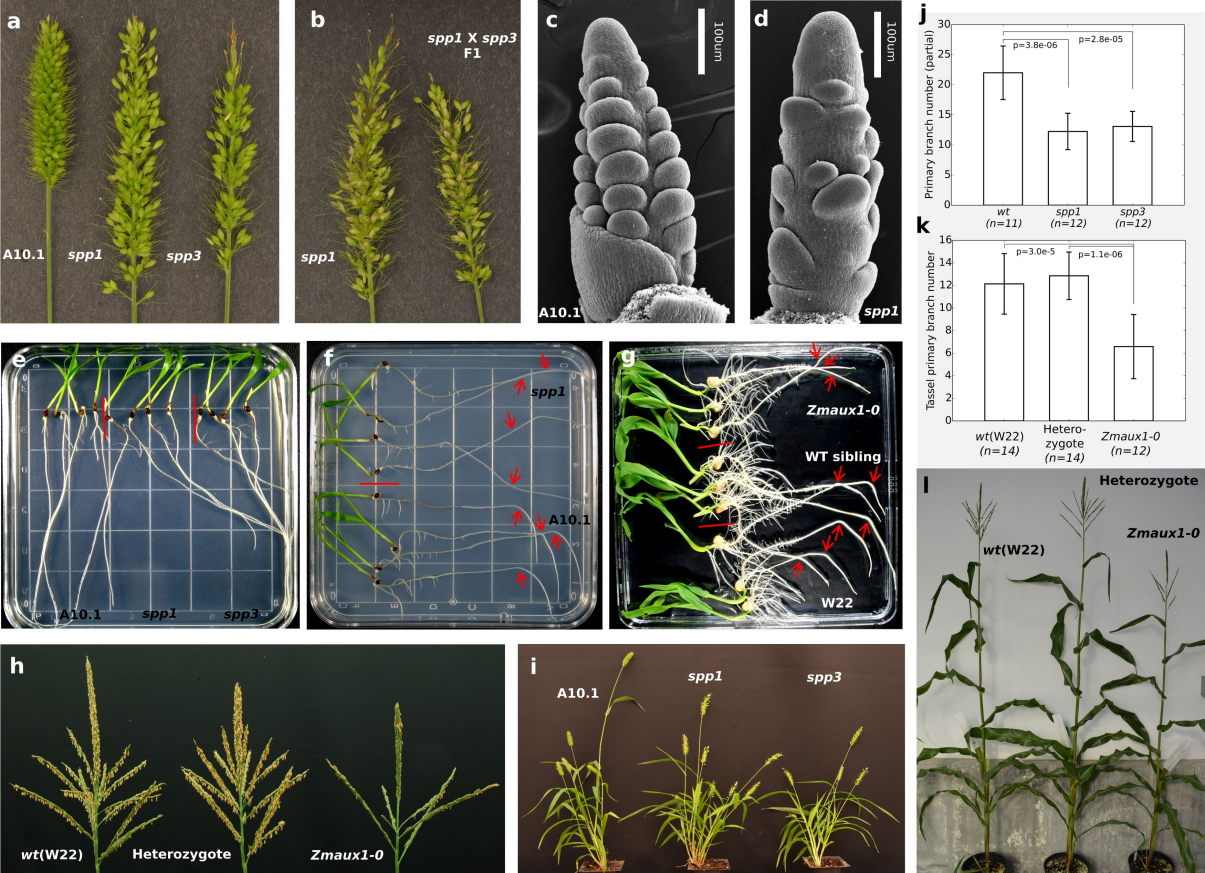
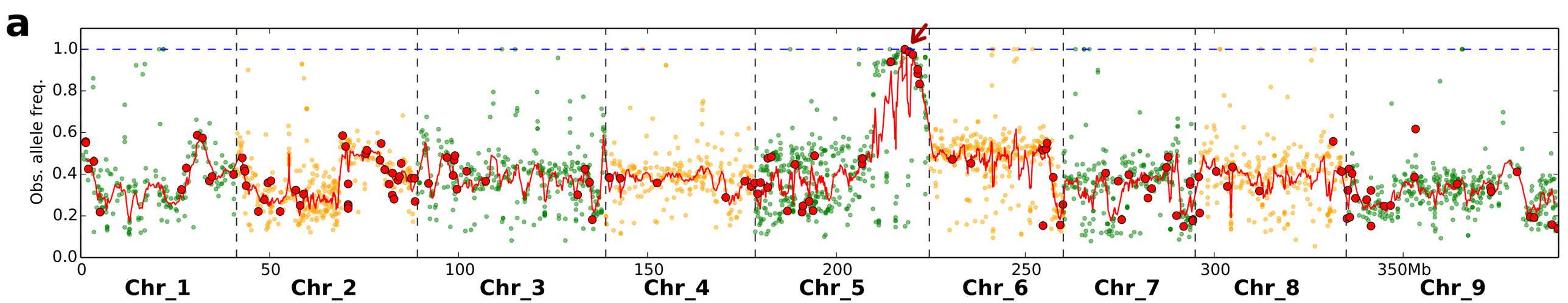
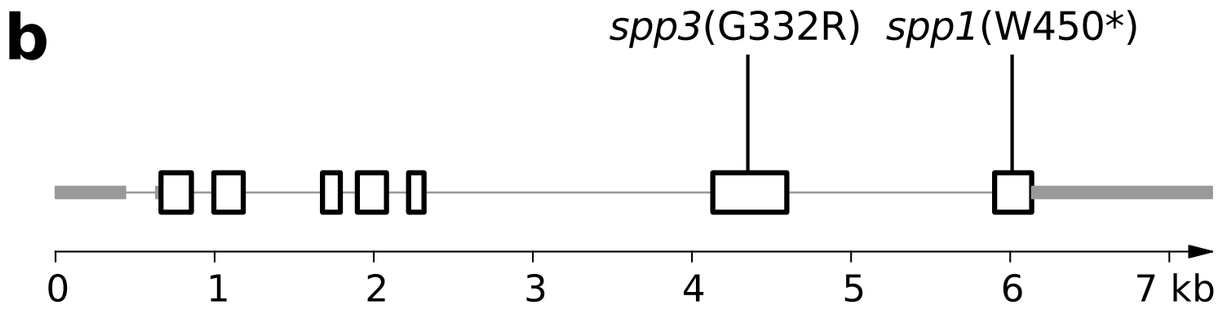
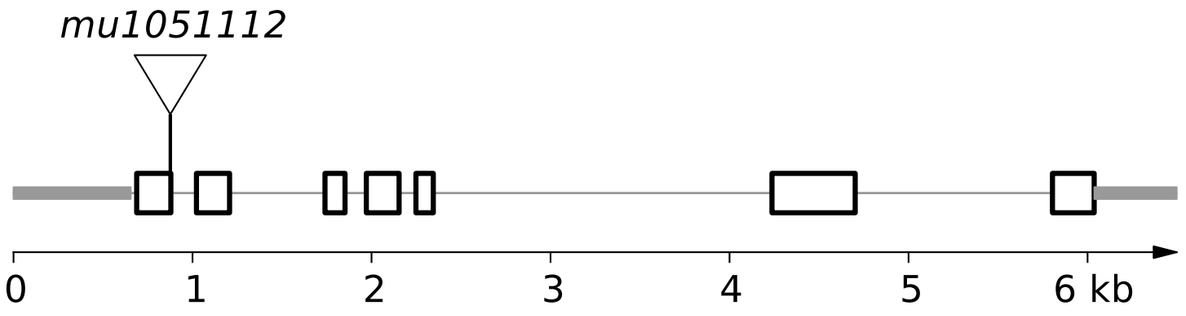


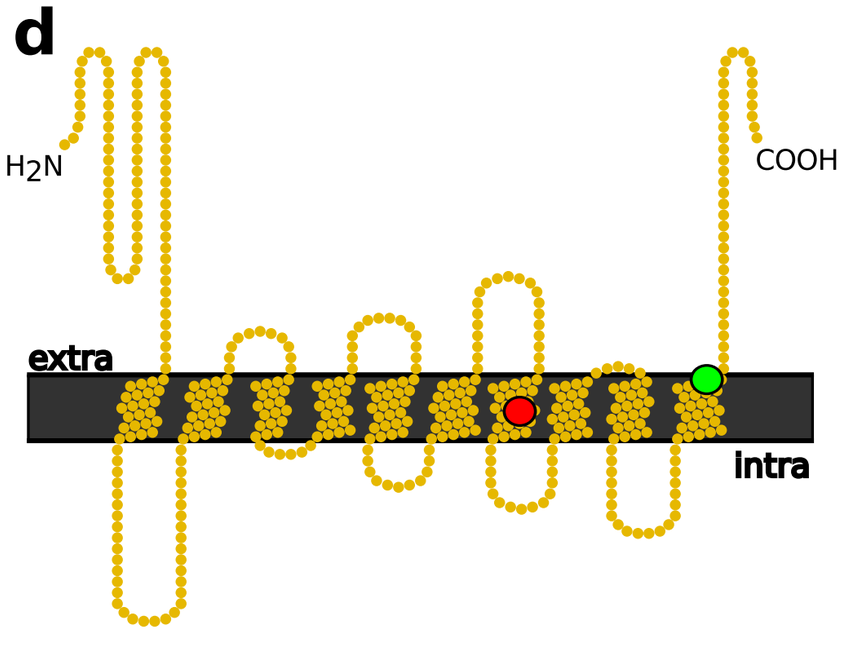
Figure 2. BSA mapping and variation of *SvAUX1* and *ZmAUX1* alleles. **a)** Bulked segregant mapping (BSA) of *spp1*, showing combined result from all five pools. The red arrow indicates the mutation in *SvAUX1* in *spp1*. **b)** Position of mutations in *SvAUX1* in *spp1* and *spp3*. **c)** Position of Mutator insertion in *ZmAUX1* in *Zmaux1-0*. **d)** Predicted transmembrane domains of *SvAUX1* protein. Red and green circles show amino acids affected by *spp1-1* and *spp1-3* mutations, respectively. **e)** Predicted transmembrane domains of *ZmAUX1* protein. Blue circle shows position of presumed protein truncation associated with *Mutator* insertion in *Zmaux1-0*. **f)** Phylogeny of *AUX1* homologs. This tree is a consensus tree of 1000 bootstraps using GTR+G+I model in RaxML. Red and green homologs denote *SvAUX1* and *ZmAUX1*, respectively. Node labels show bootstrap support percentage, and background highlight denotes five highly supported ortholog clades in grasses and one clade for homologs in *A. thaliana*. **g)** Normalized relative expression of *spp1-1* and A10.1 alleles of *SvAUX1* to reference genes Sevir.2G354200 and Sevir.9G574400 from qRT-PCR. Bar heights show mean values, error bars show standard deviations, p-values calculated from two-tailed t-test, and 3 biological replicates were made in both A10.1 and *spp1-1*.

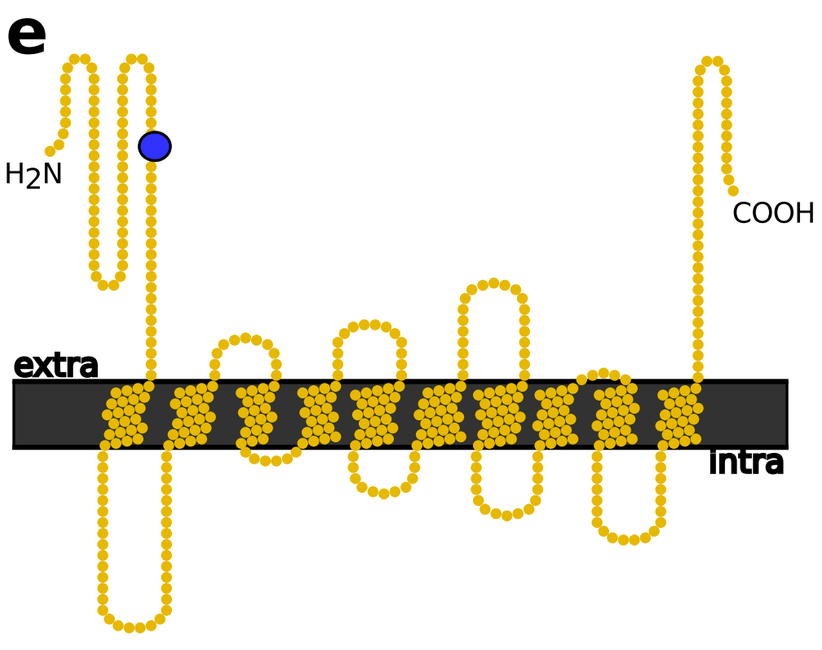

a
A10.1
spp1
spp3
b
spp1 X spp3
F1
spp1
c
100um
A10.1
d
100um
spp1
j
Primary branch number (partial)
p=3.8e-06
p=2.8e-05
wt
(n=11)
spp1
(n=12)
spp3
(n=12)
k
Tassel primary branch number
p=3.0e-5
p=1.1e-06
wt(W22)
(n=14)
Hetero-
zygote
(n=14)
Zmaux1-0
(n=12)
e
A10.1
spp1
spp3
f
spp1
A10.1
g
Zmaux1-0
WT sibling
W22
l
Heterozygote
Zmaux1-0
wt(W22)
h
wt(W22)
Heterozygote
Zmaux1-0
i
A10.1
spp1
spp3

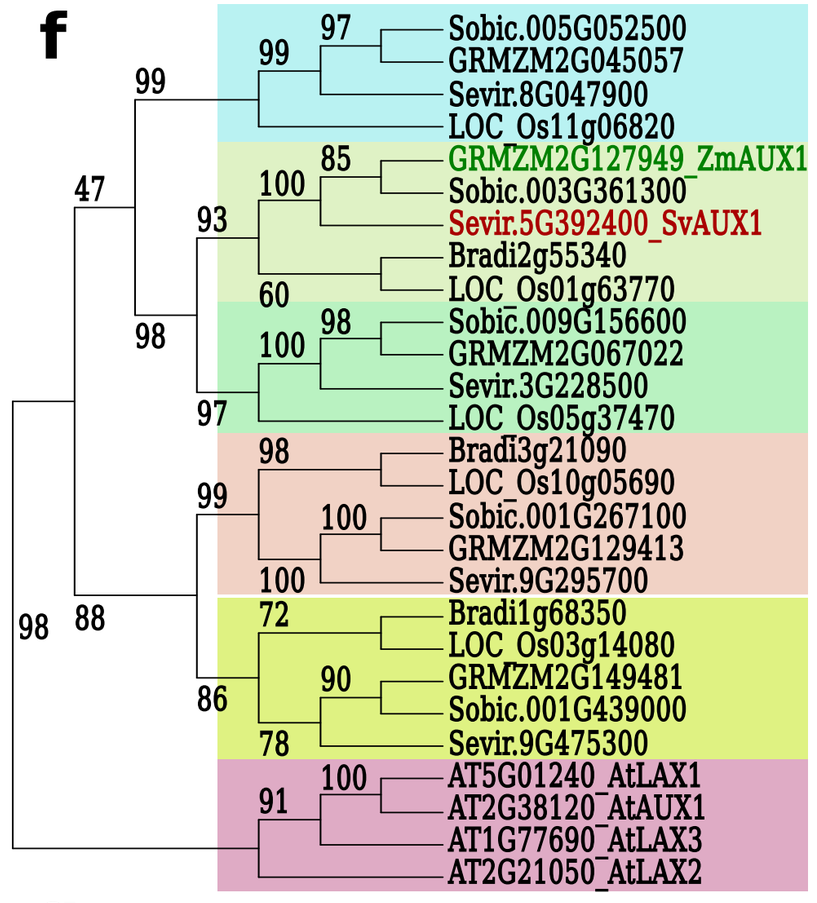

a
Obs. allele freq.
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
300
350Mb
Chr_1
Chr_2
Chr_3
Chr_4
Chr_5
Chr_6
Chr_7
Chr_8
Chr_9

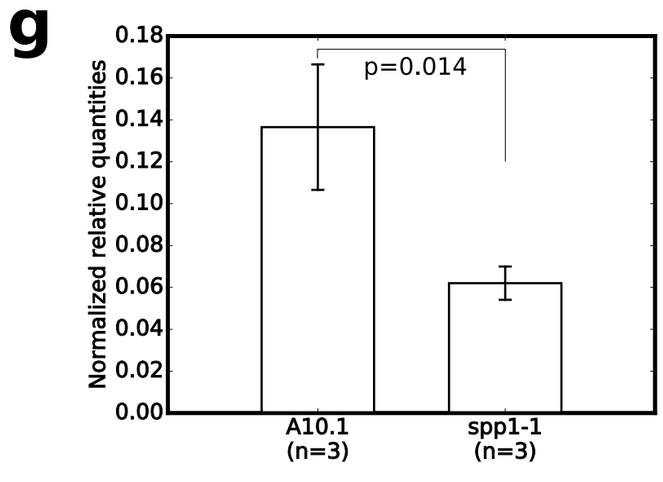

b
spp3(G332R)
spp1(W450*)
0
1
2
3
4
5
6
7 kb


c
mu1051112
0
1
2
3
4
5
6 kb


d
H2N
COOH
extra
intra


e
H2N
COOH
extra
intra


f
Sobic.005G052500
GRMZM2G045057
Sevir.8G047900
LOC_Os11g06820
GRMZM2G127949_ZmAUX1
Sobic.003G361300
Sevir.5G392400_SvAUX1
Bradi2g55340
LOC_Os01g63770
Sobic.009G156600
GRMZM2G067022
Sevir.3G228500
LOC_Os05g37470
Bradi3g21090
LOC_Os10g05690
Sobic.001G267100
GRMZM2G129413
Sevir.9G295700
Bradi1g68350
LOC_Os03g14080
GRMZM2G149481
Sobic.001G439000
Sevir.9G475300
AT5G01240_AtLAX1
AT2G38120_AtAUX1
AT1G77690_AtLAX3
AT2G21050_AtLAX2


g
Normalized relative quantities
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
p=0.014
A10.1
(n=3)
spp1-1
(n=3)